Application of a nudging technique to thermoacoustic tomography

Xavier Bonnefond and Sébastien Marinesque

December 6, 2011

Abstract

ThermoAcoustic Tomography (TAT) is a promising, non invasive, medical imaging technique whose inverse problem can be formulated as an initial condition reconstruction. In this paper, we introduce a new algorithm originally designed to correct the state of an evolution model, the *back and forth nudging* (BFN), for the TAT inverse problem. We show that the flexibility of this algorithm enables to consider a quite general framework for TAT. The backward nudging algorithm is studied and a proof of the geometrical convergence rate of the BFN is given. A method based on Conjugate Gradient (CG) is also introduced. Finally, numerical experiments validate the theoretical results with a better BFN convergence rate for more realistic setups and a comparison is established between BFN, CG and a usual inversion method.

1 Introduction

ThermoAcoustic Tomography (TAT) is a hybrid imaging technique that uses ultrasound waves produced by a body submitted to a radiofrequency pulse, uniformly deposited throughout the body. The absorption of this initial energy causes a non-uniform thermal expansion, leading to the propagation of a pressure wave outside the body to investigate. This wave is then measured all around the body with piezoelectric transducers or, more recently, thanks to interferometry techniques.

It appears that the absorption of the initial pulse is highly related to the physiological properties of the tissue [19]. As a result, the magnitude of the

1

ultrasonic emission (*i.e.* thermoacoustic signal), which is proportional to the local energy deposition, reveals physiologically specific absorption contrast. See [38, 22, 32, 44] for an introduction to the experimental setup. Considering that the initial illumination is a Dirac distribution in time, the problem of recovering the absorptivity of the investigated body from the thermoacoustic signal is equivalent to recovering the initial condition of a Cauchy problem involving the wave equation from the knowledge of the solution on a surface surrounding the imaging object [37]. Note that the experimental constraints do not allow the acquisition surface to completely surround the tissues to investigate, so that one cannot expect a better situation than measurements on a half-sphere (as in breast cancer detection for example). Moreover, several experimentation constraints limit the number of sensors, so that the methods of reconstruction considered must be robust to the space sampling.

Many works dealing with TAT have been achieved during the last decade (see [23] for an overview, or [1, 43]), and many of the methods used in these papers strongly depend on a set of additional assumptions, among which:

- the *homogeneity of the tissues*, leading to a constant speed of sound for the pressure wave;

- the *lack of frequency-dependent attenuation*, so that the pressure wave obeys the classical, undamped, wave equation;

- the *complete data situation*, where the acquisition surface is considered to enclose the whole body.

The study of TAT in the case of an homogeneous medium without attenuation gave rise to explicit inversion formulas. These latter usually require that the acquisition set is a closed surface, even if they can be approximated in the *limited view* situation (where the surface is a half-sphere for example). Among these, filtered backprojection formulas, even though they have been very successfully implemented in several situations, seem to face some issues when the source has support partly outside the observation surface, or when the data sampling is not sufficient (see [12, 15]). Fourier's type formulas, which offer a much better numerical efficiency, apply on very specific geometries (see [24, 31, 46]) or imply some additional approximations, like an interpolation from spherical to cartesian coordinates [2, 11, 13, 14, 45, 47]. Even though some of these techniques give rise to very efficient reconstruction schemes (especially in [24]), so far it is not clear if they can be extended

2

to less restrictive acquisition geometries or to an attenuated wave with non constant speed (see [25] for an answer about the acquisition surface).

Moreover, since the work of Burgholzer *et al* (see [9]), the *time reversal* method (first suggested in [12]) has been applied with success to the TAT (see *e.g.* [16, 17]). Even though the first known results on this method required a complete data framework, recent works by Stefanov *et al* extended it to a Neumann Series method in a fairly general situation (incomplete data, known variable sound speed and external source) with very good results in [33, 39].

In this paper, we show that the framework of TAT lends itself to the application of some *data assimilation* techniques. These latter, mainly used in geophysics, aim at correcting the state of an evolutionary model by means of data, in order to obtain a good approximation of the real state (see [40] for an introduction). The particular data assimilation method proposed here is based on a *nudging* technique: given an evolution model of the state and direct observation (our data), it consists in adding, inside the model equation, a newtonian recall of the state solution to the observations (or data), which is usually called the feedback or nudging term [4]. As we shall see, from a practical point of view, this method can be successfully used to manage the usual issues of the TAT inverse problem as incomplete data, external source and variable sound speed (when given, however). So far, however, the theoretical convergence result for the nudging technique is based on a classical result about stabilization of the wave equation [26, 36], which requires somehow a geometric optics condition (see [6] or [27] for an overview). This geometric condition, whose verification is not a simple matter, is not automatically satisfied in an incomplete data framework or with a variable sound speed. Consequently the proof provided in this article only stands in a very favorable situation (3d case, constant speed and complete data), but does not depend on usual stabilization techniques and yields explicit convergence rates.

Different nudging terms can be found in the literature, and most of them have been used to assimilate data in physical oceanography as in [21, 28, 30] and more recently in [5]. Fundamental articles offer the basis of most popular techniques as Kalman filter [20], Kalman-Bucy filter [8] and Luenberger observer [29].

Since the basic nudging method showed its limits in real conditions by using the data once, Auroux and Blum proposed to extend the method by adding a resolution backward in time and iterating the process, thus defining the

back and forth nudging (BFN) algorithm. They presented this technique in [3, 4] and implemented it with computationally efficient results in comparison with traditional data assimilation methods, like optimal filters (from extended Kalman's to particle filters) or other optimal minimization techniques (variational methods), which appear costfull respectively in memory needs and computing time – see the brief introductions of Talagrand in [40, 41, 42] for further details about data assimilation and a relevant bibliography. Recently, a more general formulation of the BFN using observers has been presented by Ramdani, Tucsnak and Weiss in [35].

As inverse problems are often solved by means of variational techniques, we also introduce a least squares method based on the Conjugate Gradient (CG).

The two methods introduced are compared to the Neumann Series presented in [33], which is one of the best existing method which fits the variable speed case and partial data settings.

This paper is organized as follows. In Section 2, we introduce the TAT inverse problem in a quite general form, *i.e.* without the use of the usual assumptions, then we describe the BFN algorithm and we state the main result of this article: the convergence Theorem 3. Then Section 3 is devoted to the proof of this result. In Section 4, we describe the variational method that uses CG algorithm. Finally, we give numerical results in Section 5.

Notation

In the following, we shall use this notation:

- The open ball with center $x \in \mathbb{R}^3$ and radius $r > 0$ is denoted by $B(x, r)$;

- Its boundary, the sphere with center x and radius r, is noted $S(x, r)$;

- For every $x \in \mathbb{R}^3$ and $0 < r_1 < r_2$, we define the spherical shell of center x and radii r_1, r_2:

$$A(x, r_1, r_2) = \left\{ y \in \mathbb{R}^3 | r_1 < \|x - y\| < r_2 \right\},$$

and if $x = 0$:

$$A(r_1, r_2) = A(0, r_1, r_2);$$

4

- For any set S of $\mathbb{R}^3$ and $r_1 > 0$, we denote by $T(S, r_1)$ the set S extended to the points at a distance lower than r_1 from S:

$$T(S, r_1) = \left\{ y \in \mathbb{R}^3 | \exists x \in S, \|y - x\| < r_1 \right\}.$$

2 Presentation of the method

2.1 The general TAT problem

In the following, we will denote by $p_{\text{exact}}(x, t)$ the pressure wave resulting from the thermal expansion of the body. We will make the assumption that the measurement process of this pressure is subject to some perturbation p_{noise}, such that the actual data can be written $p_{\text{data}} := p_{\text{exact}} + p_{\text{noise}}$. In thermoacoustic tomography, the data are nothing but $\{p_{\text{data}}(x, t) | x \in \mathscr{S}\}$, where $\mathscr{S}$ is a surface surrounding the body to investigate. The pressure p_{exact} satisfies the Cauchy problem (see [37] for a detailed calculation):

$$\left| \begin{array}{ll} L p_{\text{exact}}(x, t) = E_{dep}(x, t), & (x, t) \in \mathbb{R}^3 \times [-1, \infty), \\ p_{\text{exact}}(x, -1) = 0, & x \in \mathbb{R}^3, \\ \partial_t p_{\text{exact}}(x, -1) = 0, & x \in \mathbb{R}^3, \end{array} \right.$$

where L is the differential operator governing the wave, most likely the wave operator, or a damped wave operator, and E_{dep} is the energy deposited in the body around the time $t = 0$ (so we have to consider a negative initial time, $e.g.$ $t = -1$). This energy can be written:

$$E_{dep} = e(x) \frac{\mathrm{d}j}{\mathrm{d}t}(t),$$

$$e(x) := \frac{\beta c}{c_p} I_{em}(x) \Psi(x),$$

where e is called the *normalized energy deposition function*. Here β is the *thermal expansion coefficient*, c_p is the *specific heat capacity*, c the sound speed and I_{em} is the *radiation intensity* of the initial energy pulse. All these parameters, including the time shape j of the pulse, are known and we see that the knowledge of E_{dep} is sufficient to compute the absorption density Ψ inside the body, which is our purpose.

So far, the general thermoacoustic problem can be formulated in the following way:

Let $p_{\text{exact}}(x, t)$ be solution of:

$$\left| \begin{array}{ll} Lp_{\text{exact}}(x, t) = f_0(x)\dfrac{\mathrm{d}j}{\mathrm{d}t}(t), & (x, t) \in \mathbb{R}^3 \times [-1, \infty), \\ p_{\text{exact}}(x, -1) = 0, & x \in \mathbb{R}^3, \\ \partial_t p_{\text{exact}}(x, -1) = 0, & x \in \mathbb{R}^3, \end{array} \right. \qquad (1)$$

where j is known and f_0 is the object to reconstruct, with compact support in the unit ball $B(0, 1)$. Can we compute f_0, or a good approximation of f_0, from the knowledge of $p_{\text{data}} = p_{\text{exact}} + p_{\text{noise}}$ on a surface $\mathscr{S}$ surrounding $B(0, 1)$?

2.2 Reduction to the homogeneous case

As we have seen, even though it is not possible in practice, the initial energy deposition is meant to be a Dirac pulse in time at $t = 0$, that is $\frac{\mathrm{d}j}{\mathrm{d}t} = \delta_0'$. Let q_0 be solution, which is no longer to be considered for negative time values, where it is the null solution, of the Cauchy problem:

$$\left| \begin{array}{ll} Lq_0(x, t) = f_0(x)\delta_0'(t), & (x, t) \in \mathbb{R}^3 \times \mathbb{R}_+, \\ q_0(x, 0) = 0, & x \in \mathbb{R}^3, \\ \partial_t q_0(x, 0) = 0, & x \in \mathbb{R}^3, \end{array} \right.$$

which is equivalent to:

$$\left| \begin{array}{ll} Lq_0(x, t) = 0, & (x, t) \in \mathbb{R}^3 \times \mathbb{R}_+, \\ q_0(x, 0) = f_0(x), & x \in \mathbb{R}^3, \\ \partial_t q_0(x, 0) = 0, & x \in \mathbb{R}^3, \end{array} \right. \qquad (2)$$

according to Duhamel's principle (see [10]), provided that L is a (damped) wave operator. Then p_{data} is solution of (1) if and only if $p_{\text{data}} = j *_t q_0$, which means that a deconvolution operation on the thermoacoustic signal p_{exact} leads to the knowledge of the solution q_0 of (2) on the same surface $\mathscr{S}$ surrounding the object.

In the following, we will still denote by p_{exact} the solution of (2), assuming that we know this latter for all x in $\mathscr{S}$, and all t in $\mathbb{R}_+$.

2.3 Some useful facts about the standard wave equation

In this section we give some basic properties of the following wave equation:

$$\left| \begin{array}{ll} L_0 u(x,t) = 0, & (x,t) \in \mathbb{R}^3 \times \mathbb{R}_+, \\ u(x,0) = l(x), & x \in \mathbb{R}^3, \\ \partial_t u(x,0) = h(x), & x \in \mathbb{R}^3, \end{array} \right. \tag{3}$$

where l and h are two C^∞ functions with compact support in $B(0,1)$, and $L_0 := \partial_{tt} - \Delta$. All this material can be found in [18] for example. According to Huyghens' principle in the strong sense, for every couple $(x_0, t_0) \in \mathbb{R}^3 \times \mathbb{R}_+$, the classical solution $u(x_0, t_0)$ of (3) only depends on the values of l and h on the backward characteristic cone:

$$\|x - x_0\| = |t - t_0|.$$

As a matter of fact, when the initial conditions l and h have their support in $B(0,1)$, the solution at any time $\tilde{t}$ has its support in $B(0, 1 + \tilde{t})$.

Huyghens' principle leads to the energy conservation for the solution u of (3).

Definition 1. For every time T, we call *energy* of u the quantity:

$$E_u(T) := \frac{1}{2} \left(\|u(\cdot, T)\|^2_{H^1_0(\mathbb{R}^3)} + \|\partial_t u(\cdot, T)\|^2_{L^2(\mathbb{R}^3)} \right).$$

The variation of the energy of u is estimated from the following identity:

$$\partial_t u \left(\partial_{tt} u - \Delta u \right) = 0 = \partial_t \left[\frac{\|\nabla u\|^2}{2} + \frac{(\partial_t u)^2}{2} \right] - \sum_{i=1}^{3} \partial_{x_i} (\partial_t u \, \partial_{x_i} u).$$

Since u has a compact support, integrating for $0 < t < T$ and over the whole space yields:

$$E_u(T) = E_u(0).$$

Note that if h is zero, we have:

$$E_u(0) = \frac{1}{2} \|l\|^2_{H^1_0(\mathbb{R}^3)}.$$

2.4 Nudging

Now, let us introduce the application of the BFN algorithm to the general TAT problem.

Assume that the observation surface $\mathscr{S}$ is included in the sphere $S(0,1)$. As exposed in Section 2.1, our purpose is to compute an approximation of the original object f_0 from an incomplete set of data ϕp_{data}, where p_{data} is a the solution (possibly contaminated by noise) of the wave equation (2) and where ϕ allows the knowledge of p_{data} only on the observation surface $\mathscr{S}$. Theoretically speaking, ϕ should be $\mathbb{1}_{\mathscr{S}}$, but in practice the discretization process allows us to choose a C^∞ function with compact support. Indeed, given a resolution, there exists $\varepsilon > 0$ such that any function supported in $T(\mathscr{S}, \varepsilon)$ and equal to 1 on $\mathscr{S}$ has the same discrete counterpart as $\mathbb{1}_{\mathscr{S}}$ (see Figure 1). This is the reason why we will make the following assumptions on ϕ:

i $\forall x \in \mathbb{R}^3, \quad 0 \le \phi(x) \le 1$;

ii ϕ has a compact support in $T(\mathscr{S}, \varepsilon)$;

iii $\forall x \in T(\mathscr{S}, \varepsilon/2), \quad \phi(x) = 1$.

Finally, in order to avoid some technical difficulties, we will assume that the initial object f_0 has his support in $B(0, 1-\varepsilon)$ and is such that every Cauchy problem we shall consider has a classical solution, say C^∞.

We may proceed as follows:

Definition 2 (Iterate of the BFN algorithm for TAT)**.**

1. [Forward evolution]

   From an rough estimate $f_{0,1}$ of the object to reconstruct f_0, defined as an approximate of the initial condition supported in $B(0, 1-\varepsilon)$, we first consider a solution p_1 of Problem (3) with initial conditions $l = f_{0,1}$ and $h = 0$ (as usual in TAT).

   We compute this solution until time $T = 2$, chosen so that both $p_1(\cdot, t)$ and $p_{\text{exact}}(\cdot, t)$ (just as every function with compact support in the ball $B(0, 1-\varepsilon)$) vanish on $B(0, 1+\varepsilon)$ for all $t \ge T$ (cf Huyghens' principle in Section 2.3).
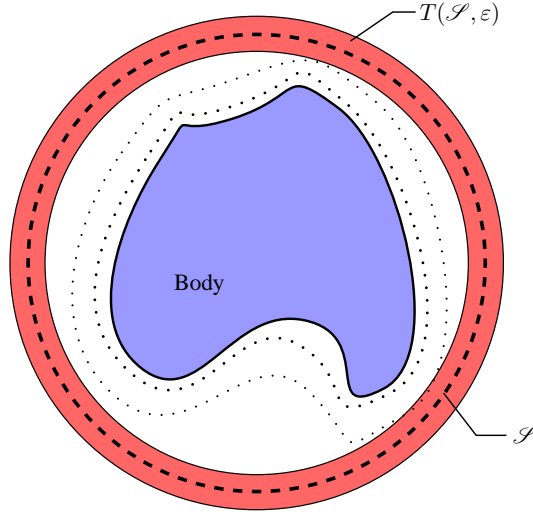
8

Figure 1: The body to investigate is surrounded by the acquisition surface $\mathscr{S}$, and generates a pressure wave.

> *Thus at this final time, in the noise-free situation, the innovation term $p_1 - p_{\text{data}}$, defined as the difference between the solution p_1 of Problem (3) and the data, is still a solution of (3) with $l = f_0 - f_{0,1}$ and $h = 0$ and has an energy matching up to:*

$$E_{p_1 - p_{\text{data}}}(T) = E_{p_1 - p_{\text{data}}}(0) = \frac{1}{2} \left\| \nabla(f_0 - f_{0,1}) \right\|_{H_0^1(\mathbb{R}^3)}^2,$$

> *since the conservation of energy for the wave equation holds in this situation.*

2. [Backward evolution]

> *Then we apply the backward nudging method to make the solution evolve back in time, with, as announced, the addition in the backward equation of a newtonian feedback, which adjusts the solution along its evolution by recalling it to the observed data.*

> *Namely, we add to L the feedback correction term $k\phi\partial_t(\cdot - p_{\text{data}})$, for some nudging parameter $k \in \mathbb{R}_+$. And we compute the backward solution starting from the final state of the forward implementation, i.e. $p_1(x,T)$ and $\partial_t p_1(x,T)$, untill the initial time $t = 0$.*

> *We obtain a corrected solution $\tilde{p}_1$ of the Cauchy problem, called back-*

ward nudging problem:

$$\left|\begin{array}{ll} L\tilde{p}_1(x,t) = k\phi(x)\partial_t(\tilde{p}_1 - p_{\text{data}})(x,t), & (x,t) \in \mathbb{R}^3 \times [0,T], \\ \tilde{p}_1(x,T) = p_1(x,T), & x \in \mathbb{R}^3, \\ \partial_t\tilde{p}_1(x,T) = \partial_t p_1(x,T), & x \in \mathbb{R}^3. \end{array}\right. \qquad (4)$$

In order to obtain an equation in forward time, let us consider the map $t \mapsto \tilde{p}_1(T-t)$, still denoted by $\tilde{p}_1$. The Cauchy problem becomes:

$$\left|\begin{array}{ll} \tilde{L}\tilde{p}_1 = -k\phi(x)\partial_t(\tilde{p}_1 + p_{\text{data}})(x,T-t), & (x,t) \in \mathbb{R}^3 \times [0,T], \\ \tilde{p}_1(x,0) = p_1(x,T), & x \in \mathbb{R}^3, \\ \partial_t\tilde{p}_1(x,0) = -\partial_t p_1(x,T), & x \in \mathbb{R}^3, \end{array}\right. \qquad (5)$$

where $\tilde{L}$ is the operator L backwards in time.

Obviously, in practice, when L is not the wave operator, there is no reason for this operator to be reversible. Nevertheless the newtonian feedback may act as a regularization term and keep the computation stable.

3. [Update of the estimate]

   Finally, after this back and forth evolution, a new estimate $\tilde{p}_1(T)$ is obtained, but in order to iterate the process we need the new estimate to be supported in $B(0, 1-\varepsilon)$, just like f_0 and $f_{0,1}$. This is simply done by defining:

   $$f_{0,2}(x) := \mathbb{1}_{B(0,1-\varepsilon)}\tilde{p}_1(x,T), \quad x \in \mathbb{R}^3.$$

The last step can be easily justified by the fact that, since f_0 is supported in $B(0, 1-\varepsilon)$:

$$\left\| f_0 - \tilde{p}_1(\cdot,T) \right\|_{H_0^1(\mathbb{R}^3)} \geq \left\| f_0 - \mathbb{1}_{B(0,1-\varepsilon)}\tilde{p}_1(\cdot,T) \right\|_{H_0^1(\mathbb{R}^3)}.$$

This scheme is then iterated, constructing a sequence of estimates $(f_{0,n})_{n\in\mathbb{N}^*}$, which may converge to f_0. The following theorem, proved in Section 3, ensures that the convergence is geometrical in $H_0^1(\mathbb{R}^3)$ under the assumptions:

(i) L is the standard wave operator $\partial_{tt} - \Delta$;

(ii) The data are noise-free, *i.e.* $p_{\text{data}} = p_{\text{exact}}$ (see Section 2.1), and the observation surface $\mathscr{S}$ is the whole sphere $S(0,1)$;

(iii) The object to reconstruct f_0 is a C^∞ function with compact support in $B(0, 1)$;

(iv) $f_{0,1}$ and $f_{0,2}$ are two consecutive estimates of f_0 obtained as described in Definition 2.

Theorem 3. Under the assumptions (i)-(iv), there exists $1 > s > 0$ only depending on f_0, ϕ and the nudging parameter $k > 0$ such that:

$$\|f_0 - f_{0,2}\|_{H_0^1(\mathbb{R}^3)} \le s \, \|f_0 - f_{0,1}\|_{H_0^1(\mathbb{R}^3)} \, . \tag{6}$$

We only consider this ideal framework for the sake of the proof, but numerical results given in Section 5 apply with more realistic assumptions and show a better convergence rate.

3 Proof of Theorem 3

This section is devoted to the proof of Theorem 3. We study the case where $L = L_0 := \partial_{tt} - \Delta$ is the standard wave operator. Let ϕ, f_0 and p_{data} be as in Theorem 3. In particular, for the sake of the proof, we make the assumptions that the data are noise-free and that, since $\mathscr{S} = S(0, 1)$, the set $T(\mathscr{S}, \varepsilon/2)$ is equal to $A(1 - \varepsilon/2, 1 + \varepsilon/2)$. We shall use the notation:

$$L_k := \partial_{tt} - \Delta + k\phi\partial_t, \quad k \in \mathbb{R}_+.$$

We first study the existence of the objects used in the BFN approximation scheme.

3.1 Solution of the backward nudging equation

The proof of the following theorem can be found in [7], Chapter 2.8.

Theorem 4. Given three C^∞-functions ϕ, l and h with compact support and $k \in \mathbb{R}_+$, $T > 0$ there exists a unique C^∞ solution p of the Cauchy problem:

$$\left|\begin{array}{lll} L_k p(x, t) = -k\phi(x)\partial_t p_{\text{data}}(x, T - t), & (x, t) \in \mathbb{R}^3 \times [0, T], \\ p(x, 0) = l(x), & x \in \mathbb{R}^3, \\ \partial_t p(x, 0) = h(x), & x \in \mathbb{R}^3. \end{array}\right. \tag{7}$$

Moreover, for every $(x_0, t_0) \in \mathbb{R}^3 \times \mathbb{R}_+$, $p(x_0, t_0)$ only depends on the values of l, h, ϕ and $p_{\text{data}}(\cdot, T - \cdot)$ in the backward characteristic cone:

$$\|x - x_0\| \leq |t - t_0|.$$

In particular, this means that for every sketch $f_{0,1}$ with compact support and of class C^∞, the solution $\tilde{p}_1$ of the backward nudging equation (7) exists, and for all $t \in [0, T]$ the function $\tilde{p}_1(\cdot, t)$ has a compact support. This property will be helpful to apply an energy method.

Now, it is easy to show that if $\tilde{p}_1$ is solution of (7) with initial conditions $l(x) = \tilde{p}_1(x, 0) = p_1(x, T)$ and $h(x) = \partial_t \tilde{p}_1(x, 0) = -\partial_t p_1(x, T)$, then the function $u := \tilde{p}_1 - p_{\text{data}}(\cdot, T - \cdot)$ is solution of the Cauchy problem:

$$\left|\begin{array}{ll} L_k u(x, t) = 0, & (x, t) \in \mathbb{R}^3 \times [0, T], \\ u(x, 0) = p_1(x, T) - p_{\text{data}}(x, T), & x \in \mathbb{R}^3, \\ \partial_t u(x, 0) = -\partial_t(p_1 - p_{\text{data}})(x, T), & x \in \mathbb{R}^3, \end{array}\right. \tag{8}$$

and satisfies:

$$E_u(0) = E_{p_1 - p_{\text{data}}}(T) = \frac{1}{2} \|f_{0,1} - f_0\|_{H^1(\mathbb{R}^3)}^2.$$

Since, from the definition of $f_{0,2}$:

$$E_u(T) \geq \frac{1}{2} \|\tilde{p}_1(\cdot, T) - f_0\|_{H^1(\mathbb{R}^3)}^2 \geq \frac{1}{2} \|f_{0,2} - f_0\|_{H^1(\mathbb{R}^3)}^2,$$

we only need to prove that:

$$E_u(T) \leq s E_u(0), \quad \text{for some } 0 < s < 1,$$

to complete the proof of Theorem 3. This is the purpose of the next section.

3.2 Energy inequality

We have the following identity for every solution u of (7):

$$\partial_t u L_k u = 0 = \partial_t \left[\frac{\|\nabla u\|^2}{2} + \frac{(\partial_t u)^2}{2} \right] + k\phi(\partial_t u)^2 - \sum_{i=1}^{3} \partial_{x_i}(\partial_t u \, \partial_{x_i} u).$$

Recall that u has a compact support for every t, thus, integrating the latter equality over the whole space and for $0 < t < T$, we get:

$$E_u(T) - E_u(0) = -k \int_0^T \int_{\mathbb{R}^3} \phi(\partial_t u)^2. \tag{9}$$

Consequently, the loss of energy of u is exactly the amount of its 'energy' passing through the support of ϕ during time T.

In [12], in order to invert the spherical Radon transform, Finch *et al* proved the following *trace identity*, which will be usefull to estimate (9):

Theorem 5. Suppose $h_i \in C_0^\infty(B(0, \rho))$ and u_i is the solution of (3) for $l = 0$ and $h = h_i$, $i = 1, 2$. Then we have the identity

$$\frac{1}{2} \int_{\mathbb{R}^3} h_1 h_2 = \frac{-1}{\rho} \int_0^\infty \int_{S(0,\rho)} t u_1 \partial_{tt} u_2.$$

From this we will obtain a key estimate of the 'energy' passing through the sphere $S(0, \rho)$. Indeed, if u is solution of (3) with $h = 0$, taking $h_1 = f$ and $h_2 = \Delta f$ in the theorem yields:

$$\frac{1}{2} \int_{\mathbb{R}^3} f \Delta f = \frac{-1}{\rho} \int_0^\infty \int_{S(0,\rho)} t (\partial_t u)^2,$$

which gives:

$$\frac{1}{2} \|f\|_{H_0^1(\mathbb{R}^3)}^2 = \frac{1}{\rho} \int_0^\infty \int_{S(0,\rho)} t (\partial_t u)^2. \tag{10}$$

Let us introduce the solution u_0 of the Cauchy problem:

$$\left|\begin{array}{ll} L_0 u_0(x, t) = 0, & (x, t) \in \mathbb{R}^3 \times [0, T], \\ u_0(x, 0) = p_1(x, T) - p_{\text{data}}(x, T), & x \in \mathbb{R}^3, \\ u_{0t}(x, 0) = -\partial_t p_1(x, T) + \partial_t p_{\text{data}}(x, 0), & x \in \mathbb{R}^3, \end{array}\right.$$

which is nothing but $(p_1 - p_{\text{data}})(\cdot, T - \cdot)$.

Then $u = u_0 + v$ where v is solution of:

$$\left|\begin{array}{ll} L_0 v(x, t) = -k\phi \partial_t (u_0 + v)(x, t), & (x, t) \in \mathbb{R}^3 \times [0, T], \\ v(x, 0) = \partial_t v(x, 0) = 0, & x \in \mathbb{R}^3. \end{array}\right.$$

The calculation of the energy of v yields:

$$E_v(T) = -k \int_0^T \int_{\mathbb{R}^3} \phi \partial_t u_0 \partial_t v - k \int_0^T \int_{\mathbb{R}^3} \phi (\partial_t v)^2.$$

Thanks to the Cauchy-Schwarz inequality, one has:

$$E_v(T) \le k \left\| \sqrt{\phi} \partial_t u_0 \right\|_{L^2([0,T] \times \mathbb{R}^3)} H(T) - kH(T)^2, \tag{11}$$

13

where:
$$H(T) := \left\| \sqrt{\phi} \partial_t v \right\|_{L^2([0,T] \times \mathbb{R}^3)}.$$

Moreover, we have:
$$E_v(T) \geq \int_{\mathbb{R}^3} (\partial_t v)^2 (\cdot, T) \geq \int_{\mathbb{R}^3} \phi(\partial_t v)^2(\cdot, T) = \frac{\mathrm{d}}{\mathrm{d}T} \left(H^2 \right)(T).$$

So that, dividing by $H(T)$ in (11) we find:
$$\frac{\mathrm{d}}{\mathrm{d}T} H(T) + kH(T) \leq k \left\| \sqrt{\phi} \partial_t u_0 \right\|_{L^2([0,T] \times \mathbb{R}^3)}.$$

A classical calculation shows that this differential inequation yields:
$$H(T) \leq \int_0^T k \left\| \sqrt{\phi} \partial_t u_0 \right\|_{L^2([0,s] \times \mathbb{R}^3)} e^{k(s-T)} \, \mathrm{d}s,$$

and since the map $s \mapsto \left\| \sqrt{\phi} \partial_t u_0 \right\|_{L^2([0,s] \times \mathbb{R}^3)}$ is increasing with s, we finally get:
$$H(T) \leq \left\| \sqrt{\phi} \partial_t u_0 \right\|_{L^2([0,T] \times \mathbb{R}^3)} (1 - e^{-kT}).$$

This inequality expresses the fact that despite the feedback part of Equation (7), its solution u keeps a certain amount of energy on the support of ϕ, which is proportional to the same energy for u_0. Indeed:
$$\begin{aligned}
\left\| \sqrt{\phi} \partial_t u \right\|_{L^2([0,T] \times \mathbb{R}^3)} &\geq \left\| \sqrt{\phi} \partial_t u_0 \right\|_{L^2([0,T] \times \mathbb{R}^3)} - H(T) \\
&\geq \left\| \sqrt{\phi} \partial_t u_0 \right\|_{L^2([0,T] \times \mathbb{R}^3)} e^{-kT},
\end{aligned}$$

thus:
$$\int_0^T \int_{\mathbb{R}^3} \phi(\partial_t u)^2 \geq \underbrace{\left(\int_0^T \int_{\mathbb{R}^3} \phi(\partial_t u_0)^2 \right)}_{A} e^{-2kT}. \tag{12}$$

The last step of this proof is the estimation of A. We have:
$$A \geq \int_{A(1-\varepsilon/2, 1+\varepsilon/2)} \int_0^T (\partial_t u_0)^2 = \int_{A(1-\varepsilon/2, 1+\varepsilon/2)} \int_0^T (\partial_t p_1 - \partial_t p_{\mathrm{data}})^2 (T-t),$$

so that:
$$A \geq \int_{A(1-\varepsilon/2, 1+\varepsilon/2)} \int_0^T (\partial_t p_1 - \partial_t p_{\mathrm{data}})^2 = \int_{1-\varepsilon/2}^{1+\varepsilon/2} \int_0^T \int_{S(0,\rho)} (\partial_t p_1 - \partial_t p_{\mathrm{data}})^2.$$

14

Since $p_1 - p_{\text{data}}$ vanishes on $A(1 - \varepsilon/2, 1 + \varepsilon/2)$ for every $t > T$:

$$A \geq \frac{1}{T} \int_{1-\varepsilon/2}^{1+\varepsilon/2} \int_0^\infty \int_{S(0,\rho)} t(\partial_t p_1 - \partial_t p_{\text{data}})^2.$$

Finally, applying (10) to $p_1 - p_{\text{data}}$, one has:

$$A \geq \frac{1}{T} \int_{1-\varepsilon/2}^{1+\varepsilon/2} \rho \frac{1}{2} \|f - f_{0,1}\|_{H_0^1(\mathbb{R}^3)}^2 \, \mathrm{d}\rho,$$

which yields:

$$A \geq \frac{\varepsilon}{T} E_u(0). \tag{13}$$

Combining the above estimate with (9) and (12), we get:

$$E_u(T) \leq (1 - \frac{k\varepsilon}{T} e^{-2kT}) E_u(0),$$

which was the inequality to prove, with $s := 1 - \frac{k\varepsilon}{T} e^{-2kT} < 1$.

Remark 6. *Since we know that $T = 2$, there exists an optimal choice for the nudging parameter k which leads to $s = 1 - \frac{\varepsilon}{8e}$. Even though this theoretical convergence rate does not seem satisfactory, we shall see in Section 5 that the actual convergence is much faster in practice.*

4 Variational approach

We deal now with a variational formulation of the inverse TAT problem, that leads to some new reconstruction techniques.

In this section, we consider the reconstruction of an object $f_0 \in H_0^1(B(0,1))$ from a set of observations $p_{\text{data}} = p_{\text{exact}} + p_{\text{noise}}$. Here, $p_{\text{exact}} := \mathscr{W} f_0$ where the linear operator $\mathscr{W}$ maps an initial condition f to its related solution of the Cauchy problem (3) with a null derivative initial condition $h = 0$, restricted to a non-empty, open observation set $A_\varepsilon \subset A(1 - \varepsilon, 1 + \varepsilon)$. Namely, $\mathscr{W}$ can be written $\mathscr{W} = \Phi W$, with:

$$W: \quad H_0^1(B(0,1)) \quad \longrightarrow \quad C^0([0,T]; H_0^1(B(0,R)))$$
$$f \quad \longmapsto \quad Wf,$$

where Wf is the weak solution of (3) with the initial conditions $l = f$ and $h = 0$. Even though Wf is defined on the whole space $\mathbb{R}^3$, we can

choose $R > 0$ large enough for any solution of (3) to keep a null trace on $S(0, R)$, such that the restriction of Wf to $B(0, R)$ is in $H_0^1(B(0, R))$ for every $t \in [0, T]$. The operator Φ is the restriction operator:

$$\Phi : \quad L^2([0, T] \times \mathbb{R}^3) \quad \longrightarrow \quad L^2([0, T] \times A_\varepsilon)$$
$$g \quad \longmapsto \quad g_{|A_\varepsilon}.$$

Our purpose is to solve the inverse problem:

$$\mathscr{W} f = p_{\text{data}},$$

by means of the following minimization problem:

$$(\mathscr{P}_\alpha) \left| \begin{array}{ll} \text{Minimize} & J_\alpha(f) \\ \text{s.t.} & f \in H_0^1(B(0, 1)). \end{array} \right.$$

Here, J_α is the Tikhonov functional:

$$J_\alpha : \quad H_0^1(B(0, 1)) \quad \longrightarrow \quad \mathbb{R}_+$$
$$f \quad \longmapsto \quad \frac{1}{2} \int_0^T \|\mathscr{W} f - p_{\text{data}}\|_{L^2(A_\varepsilon)}^2 \, \mathrm{d}t + \frac{\alpha}{2} \|f\|_{L^2(B(0,1))}^2,$$

with $\alpha > 0$. Since $\mathscr{W}$ is a bounded linear operator, J_α is a strictly convex (thanks to the regularization parameter α), differentiable and coercive functional, and any solution of $(\mathscr{P}_\alpha)$ is characterized by:

$$\nabla J_\alpha(f) = \mathscr{W}^*(\mathscr{W} f - p_{\text{data}}) + \alpha f = 0.$$

For any $\alpha > 0$, the above equation has a unique solution:

$$f_\alpha = (\mathscr{W}^* \mathscr{W} + \alpha I)^{-1} \mathscr{W}^* p_{\text{data}},$$

so that Problem $(\mathscr{P}_\alpha)$ is well-posed. We have chosen to solve this latter by means of the conjugate gradient method, which requires, at each iteration, the computation of ∇J_α.

Now, one has, for every $\psi \in C_0^\infty(B(0, 1))$ and $f \in H_1^0(B(0, 1))$:

$$\begin{aligned} \langle \nabla J_\alpha(f), \psi \rangle_{L^2(B(0,1))} &= \langle \mathscr{W}^*(\mathscr{W} f - p_{\text{data}}), \psi \rangle + \alpha \langle f, \psi \rangle \\ &= \underbrace{\langle \mathscr{W} f - p_{\text{data}}, \mathscr{W} \psi \rangle}_{G_{f,\psi}} + \alpha \langle f, \psi \rangle, \end{aligned}$$

with:

$$
\begin{aligned}
G_{f,\psi} &= \int_0^T \int_{A_\varepsilon} (Wf - p_{\text{data}})W\psi \\
&= \int_0^T \int_{B(0,R)} \mathbb{1}_{A_\varepsilon}(Wf - p_{\text{data}})W\psi.
\end{aligned}
$$

Considering the weak solution u^* of the adjoint equation:

$$
\left|
\begin{array}{l}
\partial_{tt} u^*(x,t) - \Delta u^*(x,t) = \mathbb{1}_{A_\varepsilon}(Wf - p_{\text{data}}), \\
u^*(x,T) = 0, \\
\partial_t u^*(x,T) = 0,
\end{array}
\right.
\tag{14}
$$

and since $W\psi \in C_0^\infty(B(0,R))$, we have:

$$
G_{f,\psi} = \int_0^T \int_{B(0,R)} \partial_{tt} u^* W\psi + \int_0^T \int_{B(0,R)} \langle \nabla u^*, \nabla W\psi \rangle.
$$

Since $B(0,R)$ has been chosen large enough, applying Green's formula and integrating by parts yield:

$$
\begin{aligned}
G_{f,\psi} &= \int_0^T \int_{B(0,R)} u^* LW\psi - \int_{B(0,R)} \partial_t u^*(x,0)W\psi(0)\,\mathrm{d}x \\
&= \langle -\partial_t u^*(\cdot,0), \psi \rangle.
\end{aligned}
$$

Finally:

$$
\nabla J_\alpha(f) = -\partial_t u^*(\cdot,0) + \alpha f.
$$

This calculation ensures that each iteration of the conjugate gradient algorithm necessitates successively the computations of Wf and of its adjoint state u^*, whose algorithmic complexity is comparable to one iteration of the BFN algorithm. This should be noted that the computation of the adjoint state requires the storage of $Wf - p_{\text{data}}$ on A_ε, which is not required in the BFN algorithm.

Note that even if we have formulated the calculation of ∇J_α in a continuous framework, its discrete counterpart leads to similar considerations. Obviously, the behaviour of the method strongly depends on the scheme chosen to compute solutions of (3).

5   Numerical implementation

In order to test the numerical behavior of the BFN method, we have implemented it in several situations. In particular, we have considered the case where the data are available only on a half-sphere, so that the illustration fits the practical experimental set-up.

5.1   Implementation framework

We now describe the numerical context chosen to compare Back and Forth Nudging (BFN) and the Conjugate Gradient (CG) techniques to a reasonable algorithm, namely the Neumann Series (NS) (since it is shown in [33] that the use of NS significantly improves the Time Reversal (TR) algorithm, we mainly chose to consider NS results).

We work now in a 2D framework for numerical costs reasons. The object to reconstruct is in $[-0.5, 0.5]^2$ gridded with 256^2 pixels, such that the space step is $\delta x = 1/256$.

Sensors are located on the circle $S(0, \sqrt{2}/2)$ surrounding the object, on which various distributions are considered. If no additional information is given, 800 detectors are used when the data are complete. Nevertheless some reconstructions were computed with less detectors in order to bring to light the robustness of the techniques. When a multiplicative white Gaussian noise is added to the data, their level is set to 15%.

The final time $T = \sqrt{2}$ is then taken minimal in order to leave waves traveling to the sensors when sound speed is constant with value 1. Note that the four studied methods can significantly be improved by increasing T (*cf.* [33]). We use the domain $[-0.5 - \sqrt{2} - \varepsilon; 0.5 + \sqrt{2} + \varepsilon]^2$ for the computation, where ε represents a few pixels, which is large enough to avoid reflection effects on its boundary.

The model is computed by means of the classical finite difference time domain (FDTD) method (explicit Euler scheme and classical five points discretization of the Laplace operator $\Delta_{\delta x}$).

Figure 2 represents the 2-dimensional objects to reconstruct, latter denoted by f_0. Squares able to give rise to typical phenomenons of observation of wave-like systems while the second object, known as the Shepp-Logan phan-

tom, allows good analysis and comparisons of the ability of the methods. The third object has been considered for combining sharp contours with smoother areas. Each is implemented as a 256 by 256 matrix.
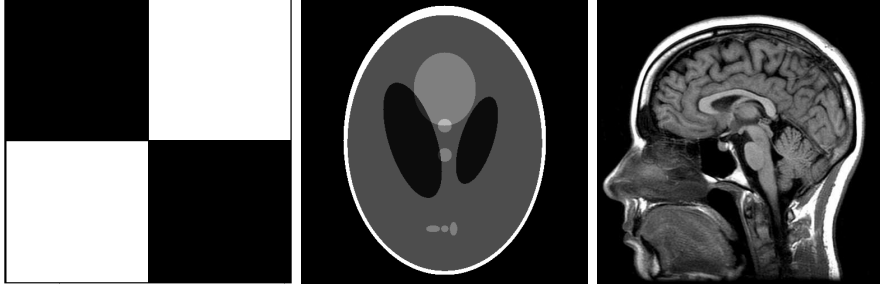

Figure 2: The initial objects f_0: squares, Shepp-Logan and skull.

Four sound speeds are used. First a constant one $c = 1$, then three other one, variable, which are chosen as in [33], respectively defined as:

$$
\begin{aligned}
c(x,y) &= 1 + 0.2\sin(2\pi x) + 0.1\cos(2\pi y), \\
c(x,y) &= \frac{9(x^2 + y^2)}{1 + 9(x^2 + y^2)} + \exp(-90(x^2 + y^2)) \\
&\quad -0.4\exp(-10(3\sqrt{x^2 + y^2} - 2)^2), \\
c(x,y) &= 1.25 + \sin(2\pi x)\cos(2\pi y).
\end{aligned}
\tag{15}
$$

They are plotted on figure 3 and abbreviated respectively NTS, TS1 and TS2, as the first is a Non-Trapping Speed while the two other are Trapping Speeds. No cutoff is used to smooth the transition between variable internal speed and constant external speed (always being 1).
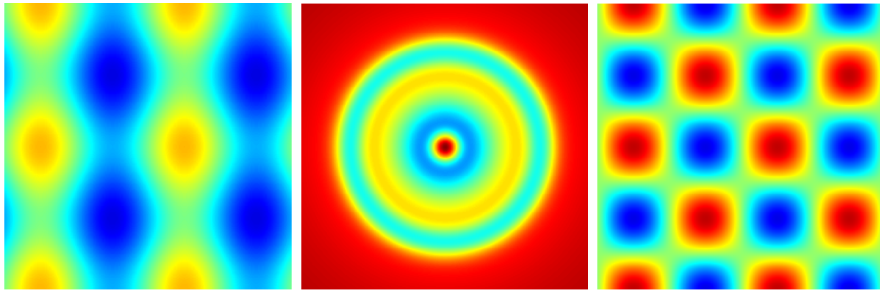

Figure 3: The different variable speeds: NTS, TS1 and TS2 (from equations (15)).

Concerning the parameters of the algorithms, the regularization parameter α in CG method equals 0. Other values were tested that stabilized the algorithm but always blur the reconstructions so that we get higher errors than without regularization. The observed convergence rates of CG and NS methods are fairly high, so that only 10 iterations yield a good approximation of the best achievable solutions. With regard to the two nudging coefficients, while we can theoretically affirm that the larger it is, the faster the solution tends to the observations (thanks to the energy equality (9)), the discretization of the time-derivative of the solution imposes a coefficient k not larger than $1/\delta t$. In order to obtain the best convergence and to make the choice of k less arbitrary (and fitting with the limits of the numerical method), it is defined as $0.9/\delta t$.

The following tables give the relative mean square error obtained with the best estimation obtained after at most 10 iterations and, in brackets, the corresponding number of iterations required. Note that BFN, CG and NS furnishes convex error plots with a minimum.

Finally, as it has been suggested in [34], we also have used an artificial numerical attenuation to offset spurious high-frequency effects and noise, with interesting consequences. Then for BFN, NS and CG techniques, the numerical model becomes:

$$\frac{u_{n+1} - 2u_n + u_{n-1}}{\delta t^2} = \Delta_{\delta x} u_n + \delta x^\varepsilon \Delta_{\delta x} \frac{u_n - u_{n-1}}{\delta t},$$

where the value for ε is empirically set and an optimal value appears for each method, which are approximately 2 for BFN and 1.7 for CG and NS in a complete data situation.

The attenuation being only introduced in a numerical improvement view, it is set to attenuate both back and forth solutions, should the case arise, and when the data are corrupted by noise.

5.2 Implementation results

The first results are given in Table 1 in a noiseless complete data situation. One sees that variable speeds can have various effects, depending on the method we use. NS offers the best reconstruction, then BFN and finally CG.

These techniques can be consequently improved by increasing the number of

Speed map	BFN	CG	NS
$c \equiv 1$	2.3(10)	7.1(3)	1.4(10)
NTS	3.2(10)	7.5(4)	2.3(10)
TS1	7.1(10)	16.5(5)	1.6(10)
TS2	7(10)	10.7(5)	6.6(10)

Table 1: Noiseless complete data for Shepp-Logan phantom: relative mean square error (number of iterations, limited to 10).

iterations as we can see in Table 2 for the BFN, where 100 iterations instead of 10 are computed.

$c \equiv 1$	NTS	TS1	TS2
1	1.5(100)	1.2(100)	4.1(100)

Table 2: Noiseless complete data from Shepp-Logan phantom: relative mean square error (BFN, 100 iterations).

When adding a 15% level noise in the constant speed situation, one sees on Table 3 that any method offers an interesting reconstruction: one gets less than 20% error for 15% noise thanks to the attenuating scheme.

BFN	CG	NS
17.6(10)	19.8(2)	17.1(2)

Table 3: Noisy complete data (15% noise) from Shepp-Logan phantom with the constant speed. Reconstruction with attenuating sheme: relative mean square error (number of iterations, limited to 10).

About the propagation of the singularities and their observation, we put the light on the effect of getting incomplete data (in a noiseless situation) in Fig. 4. The sensors are located on the upper half circle, that illustrate the practical case where the observed body cannot be surrounded by the observation surface.

In this situation (constant speed and noiseless data on the half circle), the three methods are compared in Fig. 5 and 6. NS is excellent, while BFN
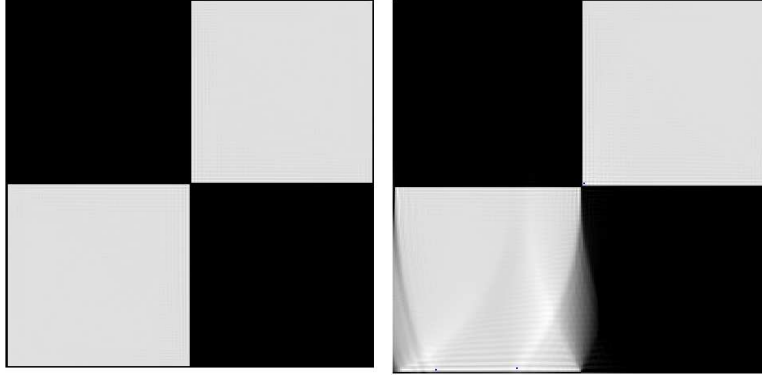
Figure 4: Reconstructions by 100 BFN iterations. Squares observed singularities. Complete (1.1% error) and half-circle (10.4%) noiseless data with constant speed.

equivalent to CG convergence in about 5 iterations, but keeps on improving the reconstruction along next iterations. As always, even if its reconstruction is the worst one, CG converges in less iterations and yields a good rough estimate faster than the others.
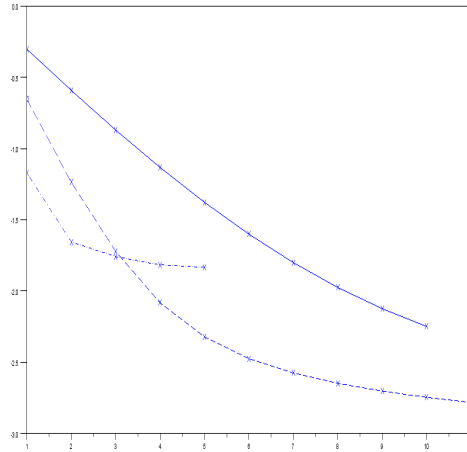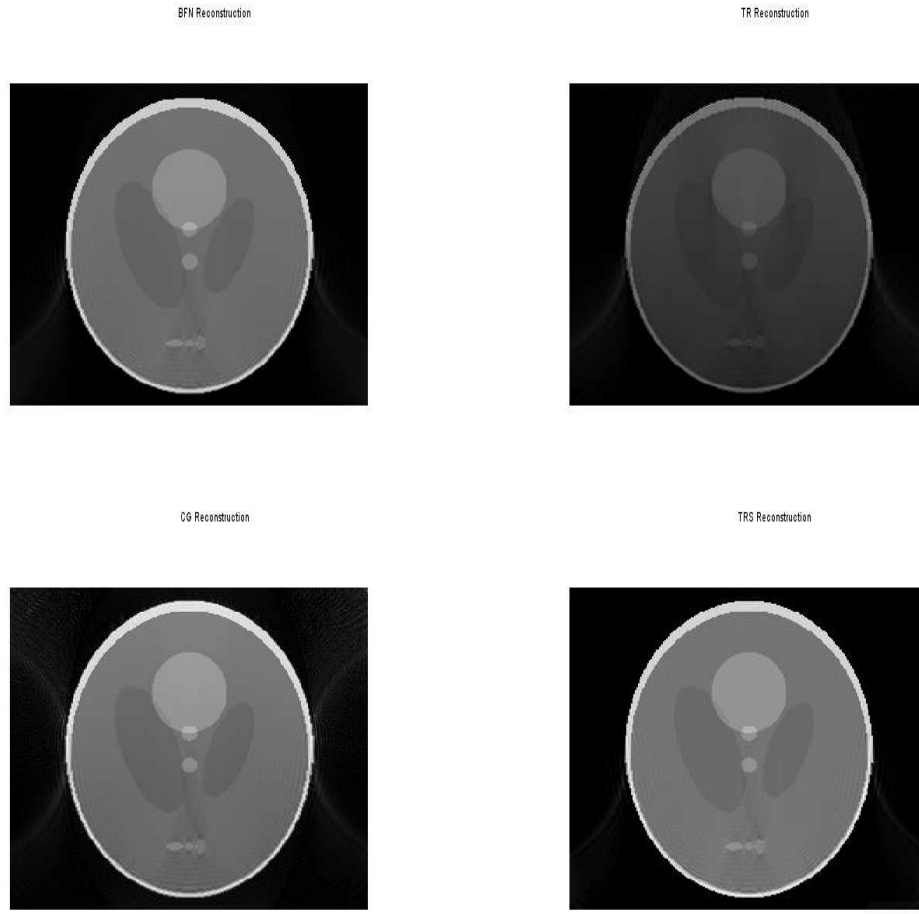

Figure 5: Incomplete noiseless data on the half circle from Shepp-Logan phantom with constant speed: relative mean square error plots (BFN in solid line, CG in dashed-dotted line and NS in dashed line).

Still testing the robustness of the methods, we reduce the number of sensors,

BFN Reconstruction

TR Reconstruction

CG Reconstruction

TRS Reconstruction

BFN	CG	TR	NS
10.6(10)	16(6)	52	6.2(10)

Figure 6: Incomplete noiseless data on the half circle from Shepp-Logan phantom with constant speed: reconstructions with BFN (top left), TR (top right), CG (bottom left) and NS (bottom right) and relative mean square errors table (error, number of iterations, limited to 10).

which are still homogeneously located on the observation circle, but only one out of σ is considered in the data. The results shown in Table 4 highlight some redundancy aspect of the data since the reconstructions have a similar quality in all situations, particularly with CG which is very robust here.

That is obvious at convergence of the algorithms, but one must consider more than 100 iterations for BFN and NS when $\sigma = 8$. Finally, if the intensity is not well recovered, forms are reconstructed with few artifacts.

Settings	BFN	CG	NS
$\sigma = 1$	2.3(10)	7.1(3)	1.4(10)
$\sigma = 2$	5.4(10)	8.9(4)	4.5(10)
$\sigma = 4$	12.1(10)	12.4(4)	11.5(10)
$\sigma = 8$	25.3(10)	15(5)	18.5(10)

Table 4: Partial noiseless data (one out of σ sensor records data) from Shepp-Logan phantom with constant speed: relative mean square errors (number of iterations, limited to 10).

Then we sum up the different unfavorable conditions from above and show the results on Fig. 7 and 8. Here data are acquired on the half circle, only one out of two sensors is considered, a 15% level noise is added and the speed is constant. The last object, a skull, has been chosen in this last experiment in order to consider a more complex structured object.
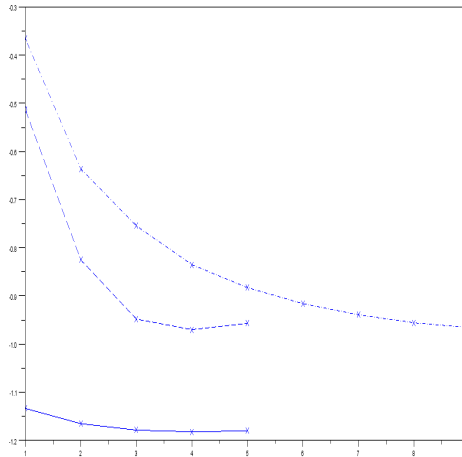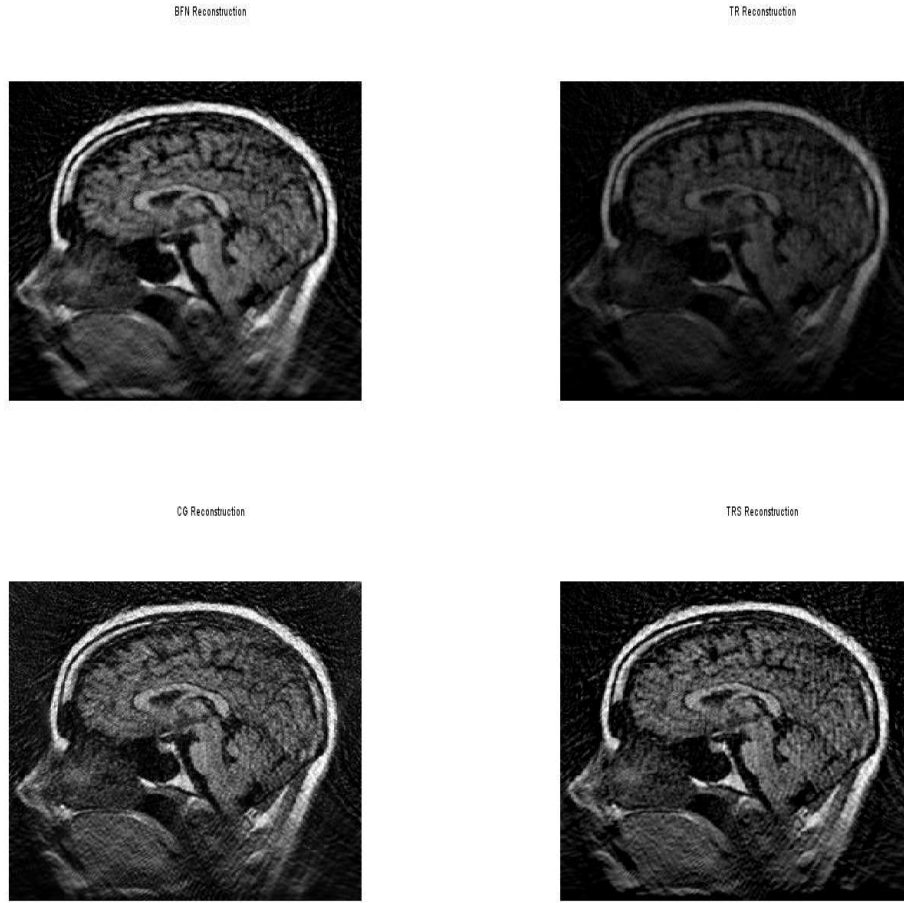

Figure 7: Incomplete partial data ($\sigma = 2$) on the half circle with 15% noise from skull phantom and with constant speed: relative mean square error plots (BFN in solid line, CG in dashed-dotted line and NS in dashed line).

Finally, even if the reconstructions obtained in this last situation are not

24

BFN Reconstruction   TR Reconstruction

CG Reconstruction   TRS Reconstruction

BFN	CG	TR	NS
30.7(4)	37.8(10)	59.9	37.9(4)

Figure 8: Incomplete 15% noise data on the half circle from skull phantom with constant speed and partial data ($\sigma = 2$): reconstructions with BFN (top left), TR (top right), CG (bottom left) and NS (bottom right) and relative mean square errors table (error, number of iterations, limited to 10).

satisfying, it may fit with very bad real conditions (where the noise level can be far greater than 15% for example) and allows us to test the robustness of the methods. In this view, Table 5 shows the errors if we consider variable speeds in a similar case but with $\sigma = 1$.

Speed map	BFN	CG	NS
$c \equiv 1$	36.4(8)	40(10)	38.7(3)
NTS	36.3(8)	40.6(10)	38.6(3)
TS1	46.3(4)	50.9(8)	51.9(3)
TS2	39.1(8)	42.3(9)	41.5(4)

Table 5: Incomplete 15% noise data on the half circle from Shepp-Logan phantom: relative mean square errors (number of iterations, limited to 10).

5.3 Implementation conclusions

CG reacts much badly to noise addition above all, but it provides good reconstructions after a couple of iterations in comparison with BFN and NS, so that it can be used to get a rough estimate for another method. Moreover CG keeps interesting when data are incomplete, not overreacting. On the whole, BFN and NS provide good and quite equivalent reconstructions with a similar cost. Finally, NS furnishes better reconstruction, but appears less robust than BFN to noise and to restriction of the number of sensors. Numerical attenuation provides a good regularization that offsets noise for NS and BFN. Notice that this artificial numerical attenuation may allow us to consider lossy medium in back and forth evolutions as the physical loss does not exceed the artificial one.

Depending on whether data are complete or not, one may choose CG to get a first rough estimate, then NS or BFN to reconstruct the object. It has to be said that not stopping iterations to 10 often improves results in noiseless and/or complete data situations, but that this number of iterations provides the best estimate in less favorable settings.

6 Conclusions

We have established the theoretical geometrical convergence of the TAT-BFN algorithm, with an explicit convergence rate in an ideal situation. In principle, this method enables to make weak assumptions in the TAT modeling (variable sound speed, damping, incomplete data, etc.), since the model is considered as a weak constraint. Even though the proof of the convergence has only been stated in an ideal framework, the numerical implementation

of the algorithm showed that the BFN method remains efficient in more general situations. Indeed this technique still yields robust reconstructions in the noisy and incomplete data situation. Even though the convergence rates of the Conjugate Gradient and of the Neumann Series methods appear to be better, the BFN provides more satisfactory asymptotic behavior and minimal relative error, along with a low numerical complexity, in quite unfavorable observation situations.

From these preliminary results, one should now focus on their generalization to more realistic wave equations. For example, one may introduce some damping in the model and study the theoretical validity and convergence of the methods. Our first attempts in this direction (mainly numerical implementations) suggest future interesting developments.
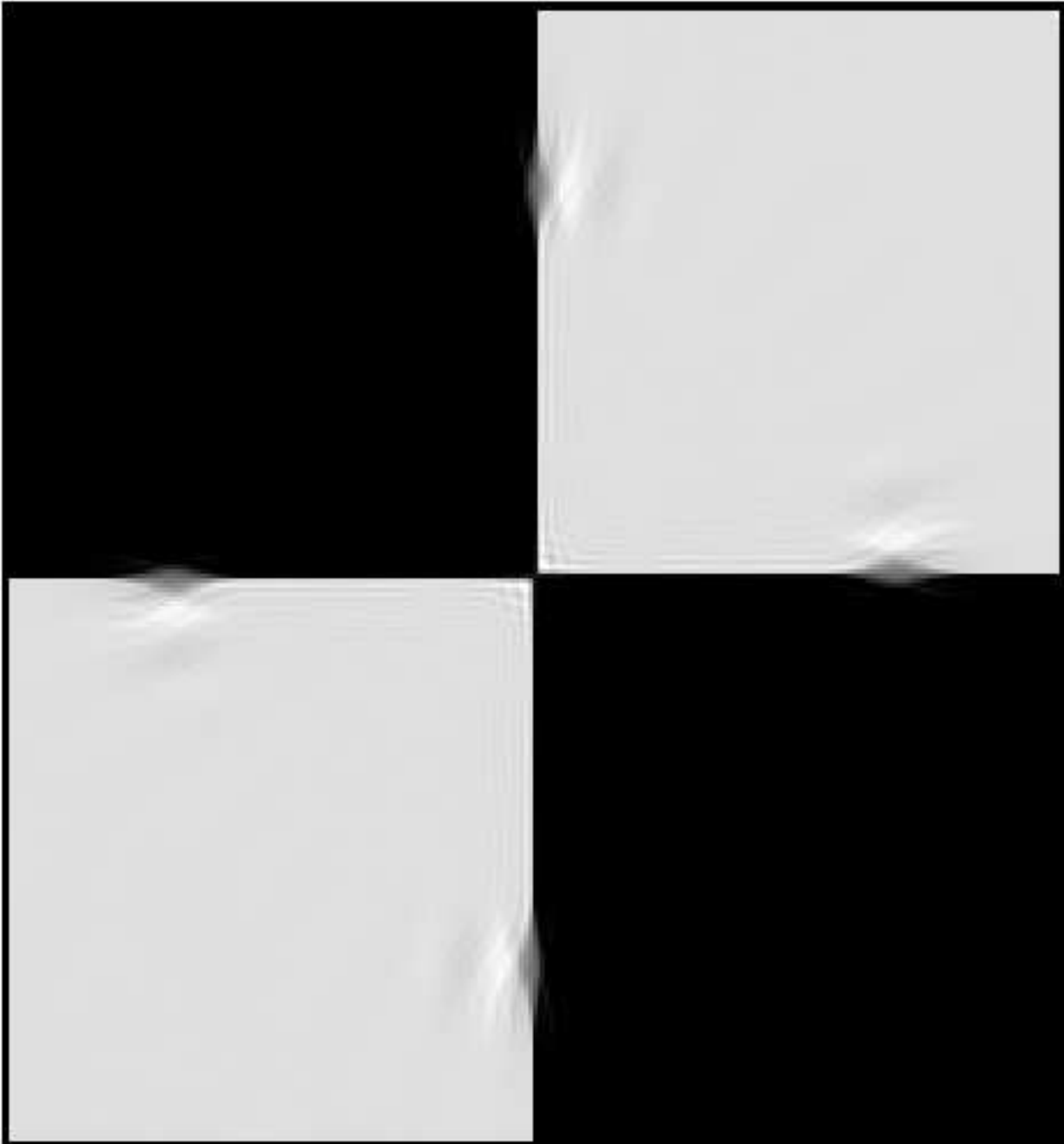
References

[1] M. Agranovsky, P. Kuchment, and L. Kunyansky. "On reconstruction formulas and algorithms for the thermoacoustic and photoacoustic tomography". In: *Photoacoustic imaging and spectroscopy. Ch. 8.* Ed. by Wang L.H. CRC Press, 2009, pp. 89–101.

[2] M.A. Anastasio et al. "Application of inverse source concepts to photoacoustic tomography". In: *Inverse Problems* 23.6 (2007), S21–S35.

[3] D. Auroux and J. Blum. "A nudging-based data assimilation method for oceanographic problems: the Back and Forth Nudging (BFN) algorithm". In: *Nonlin. Proc. Geophys.* 15 (2008), pp. 305–319.

[4] D. Auroux and J. Blum. "The Back and Forth Nudging algorithm for oceanographic data assimilation". In: *Proc. WMODA 5.* 2009, pp. 273.1–273.8.

[5] D. Auroux and S. Bonnabel. "Symmetry-based observers for some water-tank problems". In: *IEEE Trans. Automat. Contr.* 56.5 (2011), pp. 1046–1058.

[6] C. Bardos, G. Lebeau, and J. Rauch. "Sharp sufficient conditions for the observation, control and stabilization of waves from the boundary". In: *SIAM J. Contr. Optim.* 30 (1992), pp. 1024–1065.

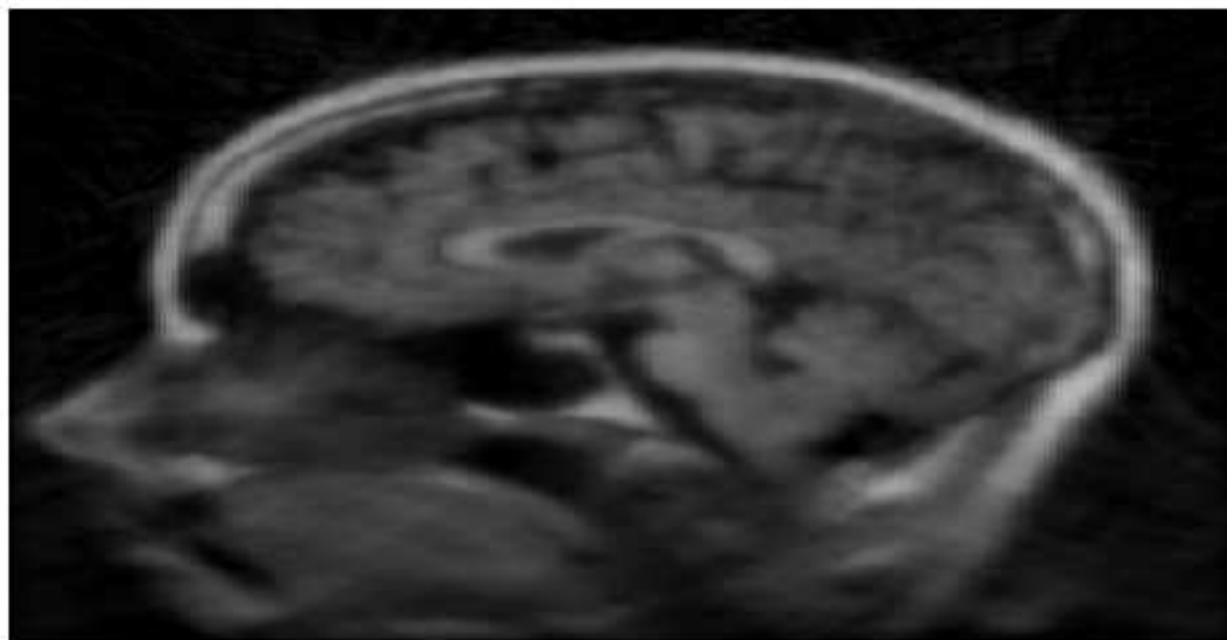[7] L. Bers, F. John, and M. Schechter. *Partial differential equations.* Interscience Publishers, 1964.

[8]   R.S. Bucy and P.D. Joseph. *Filtering for Stochastic Processes with Applications to Guidance.* Ed. by John Wiley and Sons. Providence, Rhode Island, 2nd edition, 2005.

[9]   P. Burgholzer et al. "Exact and approximative imaging methods for photoacoustic tomography using an arbitrary detection surface." In: *Phys. Rev. E* 75.4 (2007), p. 046706.

[10] R. Courant and D. Hilbert. *Meth. Math. Phys.* Ed. by Wiley. Vol. 2. 1989.

[11] B.T. Cox, S.R. Arridge, and P.C. Beard. "Photoacoustic tomography with a limited-aperture planar sensor and a reverberant cavity". In: *Inverse Problems* 23 (2007), S97–S112.

[12] D. Finch, Rakesh, and S. Patch. "Determining a function from its mean values aver a family of spheres". In: *SIAM J. Math. Anal.* 35.5 (2004), pp. 1213–1240.

[13] M. Haltmeier and O. Scherzer. "Thermoacoustic tomography and the circular Radon transform : exact inversion formula". In: *Math. Mod. Meth. Appl. Sci.* 17.4 (2007), pp. 635–655.

[14] M. Haltmeier, O. Scherzer, and G. Zangerl. "A reconstruction algorithm for photoacoustic imaging based on the nonuniform FFT". In: *IEEE Trans. Med. Imaging* 28.3 (2009), pp. 1727–1730.

[15] M. Haltmeier, T. Schuster, and O. Scherzer. "Filtered backprojection for thermoacoustic computed tomography in spherical geometry". In: *Math. Meth. Appl. Sci.* 28 (2005), pp. 1919–1937.

[16] Y. Hristova. "Time reversal in thermoacoustic tomography - an error estimate." In: *Inverse Problems* 25.055008 (2009).

[17] Y. Hristova, P. Kuchment, and L. Nguyen. "Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media". In: *Inverse Problems* 24.055006 (2008). DOI: 10.1088/0266-5611/24/5/055006.

[18] F. John. *Plane Waves and Spherical Means.* Ed. by Wiley Interscience. New York, 1955.

[19] W. Joines et al. "The measured electrical properties of normal and malignant human tissues from 50 to 900 mhz". In: *Med. Phys.* 21 (1994), pp. 547–550.

[20] R.E. Kalman. "A new approach to linear filtering and prediction problems". In: *Trans. ASME, ser. D, J. Basic Eng.* 82 (1960), pp. 35–45.

[21] T.N. Krishnamurti et al. "Physical initialisation for numerical weather prediction over the tropics". In: *Tellus* 43A (1991), pp. 53–81.

[22] R.A. Kruger, K.M. Stantz, and W.L Kiser. "Thermoacoustic CT of the breast". In: *Proc. SPIE* 4682 (2002), pp. 521–525.

[23] P. Kuchment and L. Kunyansky. "Mathematics of Photoacoustic and Thermoacoustic Tomography". In: *arXiv:0912.2022v1* (2009).

[24] L. Kunyansky. "A series solution and a fast algorithm for the inversion of the spherical mean Radon transform". In: *Inverse Problems* 23 (2007), s11–s200.

[25] L. Kunyansky. "Thermoacoustic tomography with detectors on an open curve: an efficient reconstruction algorithm". In: *Inverse Problems* 24.5 (2008).

[26] J.-L. Lions. "Exact controllability, stabilization and perturbations for distributed systems". In: *SIAM Rev.* 30.1 (1988), pp. 1–68.

[27] K. Liu. "Locally distributed control and damping for the conservative systems". In: *SIAM J. Contr. Optim.* 35.5 (1997), pp. 1574–1590.

[28] A.C. Lorenc, R.S. Bell, and B. Macpherson. "The meteorological office analysis correction data assimilation scheme". In: *Q. J. R. Meteorol. Soc.* 117 (1991), pp. 59–89.

[29] D.G. Luenberger. "Observers for Multivariable Systems". In: *IEEE Trans. Autom. Cont.* AC–11:2 (April 1966), 190–197.

[30] W.H. Lyne, R. Swinbank, and N.T. Birch. "A data assimilation experiment and the global circulation during the FGGE special observing periods". In: *Q. J. R. Meteorol. Soc.* 108 (1982), pp. 575–594.

[31] S.J. Norton and M. Linzer. "Ultrasonic reflectivity imaging in three dimensions: Exact inverse scattering solutions for plane, cylindrical and spherical apertures". In: *IEEE Trans. Med. Imaging* 28 (1981), pp. 2002–2020.

[32] G. Paltauf et al. "Photoacoustic tomography using a Mach-Zehnder interferometer as acoustic line detector". In: *Appl. Opt.* 46 (2007), pp. 3352–3358.

[33] J. Qian et al. "An efficient Neumann-series based algorithm for Thermoacoustic and Photoacoustic Tomography with variable sound speed". In: *to appear in SIAM J. Imag. Sci.* (2010).

[34] K. Ramdani, T. Takahashi, and M. Tucsnak. "Uniformly exponentially stable approximations for a class of second order evolution equations - Application to LQR problems". In: *ESAIM Control Optim. Calc. Var.* 13 (2007), pp. 503–527.

[35] K. Ramdani, M. Tucsnak, and G. Weiss. "Recovering the initial state of an infinite-dimensional system using observers". In: *Automatica* 46 (2010), pp. 1616–1625.

[36] D. L. Russel. "Controllability and stabilizability theory for linear PDE's:recent progress and open questions". In: *SIAM Rev.* 20 (1978), pp. 639–739.

[37] O. Scherzer et al. *Variational methods in imaging.* Ed. by Springer. Appl. Math. Sci. 196. New York, 2009.

[38] K.M. Stantz, R.A. Kruger, and W.L. Kiser. "Thermoacoustic CT of the breast". In: *Proc. SPIE* 4682 (2002), pp. 521–525.

[39] P. Stefanov and G. Uhlmann. "Thermoacoustic tomography with variable sound speed". In: *Inverse Problems* 25.075011 (2009).

[40] O. Talagrand. "Assimilation of Observations, an Introduction". In: *J. Meteorol. Soc. Japan* 75, No. 1B (1997), pp. 191–209.

[41] O. Talagrand and P. Courtier. "Variational assimilation of meteorological observations with the adjoint vorticity equation. I : Theory". In: *Q. J. R. Meteorol. Soc.* 113 (1987), pp. 1311–1328.

[42] O. Talagrand and P. Courtier. "Variational assimilation of meteorological observations with the adjoint vorticity equation. II : Numerical Results". In: *Q. J. R. Meteorol. Soc.* 113 (1987), pp. 1329–1347.

[43] L.H. Wang, ed. *Photoacoustic imaging and spectroscopy.* Boca Raton, FL: CRC Press, 2009.

[44] M. Xu and L.V. Wang. "Photoacoustic imaging in biomedicine". In: *Rev. Sci. Instrum.* 77.4 (2006), p. 041101.

[45] Y. Xu, D. Feng, and L. V. Wang. "Exact frequency-domain reconstruction for thermoacoustic tomography: I. Planar geometry". In: *IEEE Trans. Med. Imaging* 21.7 (2002), pp. 823–828.

[46] Y. Xu, M. Xu, and L. V. Wang. "Exact frequency-domain reconstruction for thermoacoustic tomography: II. Cylindrical geometry". In: *IEEE Trans. Med. Imaging* 21.7 (2002), pp. 829–833.

[47] Y. Xu et al. "Reconstructions in limited-view thermoacoustic tomography". In: *Med. Phys.* 31.4 (2004), pp. 724–733.

BFN Reconstruction

CG Reconstruction